I/O Performance Analysis Tool for Performance Measurement Data from Scientific Clusters

Project Description

Scientific clusters have encountered technical advances that involve increasing scales of data volumes, number of machines, and exploited parallelism in software executions. This leads to unforeseen scales of interactions in software executions between hardware components and nodes. Reasons such as details of underlying hardware and runtime information of a cluster containing large number of nodes as well as large-scale performance related logs and noises from the interactions and interferences in the executions of a cluster make diagnosis of I/O performance bottlenecks on scientific clusters more challenging.

Scientific clusters generally consist of a job scheduling engine, compute nodes for computations from an assigned job, storage nodes for a parallel file system, data transfer nodes for network accesses, and special purpose nodes for database or web services. Scientific applications are usually implemented with parallel programming framework such as Message Passing Interface (MPI). To execute an MPI-type application on a scientific cluster, an application job is submitted to the job scheduling engine. Considering the requested resources for the job and current system load, the engine assigns one or multiple nodes to the job. The resources are reclaimed when the job finishes with success or is ended accidentally due to a failure or explicit termination of the job. The log contains multiple fields providing information about job specification such as executable name and parameters, cluster assignment such as host name, and resource usages in the executions. The portions of the parallel executions from an MPI job are divided into tasks, and the tasks can be assigned to multiple nodes. Since a node can be assigned for multiple jobs, the application may experience performance degradation from interferences in shared hardware resources due to the executions of other co-located jobs. As the system loads dynamically change over time and the interferences depend on the execution characteristics of applications, this makes diagnosing I/O performance in scientific clusters more challenging.

This project is motivated by the observations that I/O performance analyses can be conducted from monitored performance measurements from scientific clusters. We aim to develop a software tool implementing I/O performance analysis on performance measurement data for resource usages. We further plan to identify an abnormal behavior in the performance of the application and computer systems. Early detection of an I/O bottleneck and performance abnormality will increase efficiency of the application.

Task Goals

The main research goal is to develop an I/O performance analysis tool for the performance measurement data from the NERSC scientific cluster. It includes following items:

- Development of a parser for the Procmon measurement data
- Development of an interactive input module for the I/O performance analysis tool
- Development of a presentation output module for the I/O performance analysis tool

Task Requirements

- Must be proficient in a programming language, such as Python.
- Good problem solving skills and communication skills
- Must be enthusiastic about exploring research literature and identifying research challenges

About the group

The Scientific Data Management (SDM) group at Lawrence Berkeley National Laboratory develops technologies and tools for efficient data access, data storage, data analysis, and management of massive scientific data sets. We are currently developing storage resource management tools, data querying technologies, in situ feature extraction algorithms, data analysis algorithms, along with software platforms for exascale data. The group also works closely with application scientists to address their data processing challenges. These tools and application development activities are backed by active research efforts on novel algorithms for emerging hardware platforms.